

**School of Information and Library Science
University of North Carolina, Chapel Hill
INLS 690–161: Digital Library Policies**

Fall 2013

Meeting Time: Friday 9:00 – 11:45, October 11 – Dec 6

Location: 304 Manning

Credits: 1.5

Instructor: Reagan Moore

Office: 22 Manning (Moore)

Phone: 919-962-9548 (Moore)

E-Mail: rwmooore [at] renci [dot] org

Office Hours: 2:00 PM on Friday, or by appointment

Course Web Site:

<https://sakai.unc.edu/portal/site/f21f40f7-49eb-4e18-bcff-1fcf24a2a778>

COURSE DESCRIPTION

Digital libraries make assertions about their digital contents, such as arrangement, descriptive metadata, access controls, completeness, and retention. We will explore the development of policies that enforce the desired digital library properties, and apply the policies to student collections managed in the LifeTime Library. Students will be able to create a collection, define policies for managing the collection, and apply the policies to their own collection.

This class will prepare students to develop and implement policies for validating digital repository trustworthiness and for management of digital collections. This includes formulation of policies that can be enacted through computer actionable rules, adapting existing rules and developing new rules. The rules will be applied in the LifeTime Library. The underlying technology is the integrated Rule-Oriented Data System (iRODS), which organizes distributed data into a sharable collection. Rules are used to automate collection administration, or enforce policies, or validate assessment criteria. Students will define policies and identify rules for particular digital management situations.

Topics covered will include: policy-based data management systems; the LifeTime Library, policies needed by digital libraries, ingestion policies, management policies, administrative policies, and assessment policies.

Students will receive accounts in the LifeTime Library, and will be able to specify rules for managing their personal digital library. Previous experience with programming will be very helpful but is not required. Knowledge of the material in INLS-461 "Information Tools" will be very helpful.

Institutions applying policy-based data management systems include:

- UNC-CH SILS LifeTime Library
- UNC-CH Carolina Digital Repository

- French National Library
- Texas Digital Libraries
- UNC-CH Genomics data grid
- NOAA National Climatic Data Center
- NASA Center for Climate Simulation

COURSE OBJECTIVES

Upon completion of this course, you should be able to:

- Define a representative set of rules appropriate to a specific data management context (e.g. shared collection, digital library, preservation environment, reference collection)
- Test rules on your LifeTime Library
- Identify ways to verify whether a rule set correctly enforces desired collection properties
- Articulate requirements for trustworthy and sustainable repositories
- Express data management policies that can be supported by computer-executable rules that control operations performed within the iRODS data grid

PREREQUISITE KNOWLEDGE REQUIRED

Knowledge of the C programming language or a scripting language is desirable, but it is not required. Sufficient information will be provided for students to define rule sets on their own.

HARDWARE AND SOFTWARE REQUIREMENTS

Students should have access to a laptop (Mac, Windows, or Unix operating system) and will access the LifeTime Library through a web browser. If you foresee any problems with this laptop requirement, you should let the instructors know as soon as possible. Students will receive an account within the SILS LifeTime Library. Students will apply rules to control their own personal data grid.

COURSE EXPECTATIONS

- Complete readings BEFORE CLASS each week. Manage your time accordingly.
- Come to class on time.
- Participate in discussions – counts as 20% of your total grade for the course.
- Demonstrate concerted effort to successfully complete all lab exercises, and submit products from the exercises that reflect this effort.
- Practice "respectful and informed ignorance." Will Rogers said, "Everybody is ignorant, only on different subjects." This class will be most effective if everyone feels comfortable asking questions, so respect the questions of others. Bring to class your own informed questions about the week's materials (i.e. be able to convey how you've tried to understand the issues and what still remains unclear to you).

Special Needs: If you feel that you may need an accommodation for a disability or have any other special need, please make an appointment to discuss this with the instructor. We will best be able to address special circumstances if we know about them early in the semester.

COURSE REQUIREMENTS

1. Adequate preparation – read required materials each week
2. Participation in class discussions – active engagement with course material, raising questions, contributing to classroom discussions
3. Class participation
4. Completion of in-class lab exercises - Instructions will be provided each week as part of the lab.
6. Final exam - This will be an in-class exam. In order to prepare, you will be provided a list of questions near the end of the semester, and the actual exam questions will be a subset of the list you have already received.

EVALUATION

- Class participation: 25%
- Completion of in-class lab exercises: 25%
- Presentation of policies for your LifeTime Library: 25%
- Final exam: 25%

Based on UNC Registrar Policy for graduate-level courses (<http://regweb.unc.edu/resources/rpm24.php>), both assignment and semester grades will be H, P, L or F. Few students will obtain an "H," which signifies an exceptionally high level of performance (higher than an "A" in an A-F systems). The following is a more detailed breakdown:

H = Superior work: complete command of subject, unusual depth, great creativity or originality

P = Satisfactory performance that meets course requirements (expected to be the median grade of all students in the course).

L = Unacceptable graduate performance: substandard in significant ways

F = Performance that is seriously deficient and unworthy of graduate credit

COURSE READINGS

Required Text:

The texts for this course will be provided as pdf files. Chapters 1-3 of the following book describes the iRODS approach for creating a computer actionable rule. We will use similar rules to control properties of the LifeTime Library.

Ward, Jewel, Michael Wan, Wayne Schroeder, Arcot Rajasekar, Antoine de Torcy, Terrell Russell, Hao Xu, and Reagan Moore. *The integrated Rule-Oriented Data System (iRODS) Micro-service Workbook*, ISBN 978-1-46646-912-9, Amazon.com. (pdf file available through the class Sakai web site)

Additional references will be available from the Sakai web site, and include:

Actionable Management Policies

Comments on Digital Libraries

ISO 16363 Metrics Evidence

LifeTime Library at UNC-CH
Policy Driven Ingestion
Towards a Theory of Digital Preservation

COURSE SCHEDULE AND TOPICS

NOTE: Most weeks of this course will follow a structure of lecture and discussion, followed by application of the policies to the LifeTime Library.

Week 1 (October 11) - Course Introduction and Preparation

We will discuss policy-based data management systems, the capabilities they provide, and how they are used to support a variety of data management applications.

Read:

- Moore, Reagan. "Towards a Theory of Digital Preservation." *International Journal of Digital Curation* 1, No. 3 (2008).
<http://www.ijdc.net/index.php/ijdc/article/viewFile/63/42>

Class topics:

- Policy Based Data Management overview
- Applications in data grids, digital libraries, persistent archives
- Properties / policies / procedures / assessment criteria
- LifeTime Library accounts

Week 2 (October 18) – Overview of LifeTime Library

We will become familiar with the LifeTime Library, log onto the system, and build a test digital library. We will explore the rules used in the LifeTime Library for integrity, authenticity, chain of custody.

Read:

- “LifeTime Library at UNC-CH”
- “Comments on Digital Libraries”

Class topics:

- LifeTime Library web browser interface
- LifeTime Library synchronization interface
- LifeTime Library default policies
- LifeTime Library rule engine

Week 3 (October 25) – Policies for Digital Libraries

We will explore the types of policies that are needed to manage a digital library. The broad categories are ingestion, management, administration, and assessment.

Read:

- “Micro-Services Workbook, chapters 1-3.

- Lab: Each student will generate a list of desired policies for each category
- We will compare lists to define a consensus on a common set of policies
- We will look for examples of application of the consensus policies
- Introduction to simple computer actionable rules

Week 4 (November 1) – Ingestion Policies

We will explore types of policies that automate the ingestion of material into a digital library. This includes automated deposition of files, extraction of metadata, creation of required data formats, and arrangement within a collection.

Read:

- “Policy driven ingestion”, Mike Conway

Class topics:

- Choosing the properties for the digital library
- Mapping properties to ingestion criteria
- Mapping ingestion criteria to policies
- Automating enforcement of ingestion policies
- Introduction to simple rule for arranging data

Week 5 (November 8) – Management Policies

We will examine policies that are required to enforce desired properties of a digital library. These management policies control the procedures used to manage the digital library. Examples include retention and disposition of material, distribution of material, access controls, and integrity constraints.

Read:

- Moore, Reagan W. "Building Preservation Environments with Data Grid Technology." *American Archivist* 69, no. 1 (2006): 139-58.
<http://www.metapress.com.libproxy.lib.unc.edu/content/176p5112w5278567/fulltext.pdf>

Class topics:

- Mapping digital library properties to management criteria
- Expressing management criteria as policies that control library procedures
- Enforcing periodic application of library procedures
- Tracking status of management procedures
- Introduction to simple rule for verifying retention period

Week 6 (November 15) – Administrative Policies

We will examine automation of tasks that are normally associated with Information Technology infrastructure. When collections become large, on the order of millions of files, automation of administrative functions is required. Examples include migration of files to new storage, generation of usage reports, management of staff accounts, and automated creation of required data formats.

Read:

- “Actionable Management Policies”

Class topics:

- Survey of standard administrative functions
- Mapping of administrative functions to policies
- Creation of a simple rule to generate usage reports
- Creation of a simple rule to version a file

Week 7 (November 22) – Assessment Policies

The properties of a digital library can be verified by policies that are run periodically. A very common assessment policy is the verification of the integrity of the digital files (replicas and checksums). Examples of other policies are related to preservation (authenticity, chain of custody, trustworthiness).

Read:

- [“ISO 16363 Metrics Evidence-100213”](#)

Class topics:

- ISO 16363 – Trustworthiness assessment criteria review
- Translation of assessment criteria to policies
- Identification of properties of production rules
- Example of a integrity verification rule

Week 8 (November 29) – Holiday

Week 9 (December 6) – Final

The exam will consist of essay questions about types of policies that can be used, and applications of policy-based data management systems.